
Rule WLM026: Highest importance service class period had few samples

Finding: A service class period at the highest performance goal importance had few Using or Delay samples for a significant percent of the RMF measurement intervals being analyzed.

Impact: This finding should be viewed as generally having a MEDIUM IMPACT or HIGH IMPACT on the performance of the Workload Manager and its ability to manage system resources to meet your performance goals.

Logic flow: This a basic finding. There are no predecessor rules.

Discussion: The Workload Manager generates control blocks to describe each service class period defined in the service policy. These control blocks describe the performance goal, the importance of the goal, the resource group (if any) to which the service class is assigned, etc.

Additionally, the Workload Manager generates control blocks for each service class period to describe the resources used by the service class period, delays to the service class period, etc.

The Workload Manager periodically examines the SRM control blocks describing each address space and acquires samples¹ describing the state of each dispatchable unit of an address space (that is, each TCB or SRB associated with the address space). The Workload Manager accumulates the samples into counters which describe the state of the address space². The samples are summarized by service class period.

The Workload Manager uses the samples in two basic ways: (1) the samples of CPU using and delays are used in computing the execution velocity of service class periods³, and (2) the samples in various categories are used to guide the actions that the Workload Manager might take as it attempts to reallocate resources to meet specified performance goals.

¹With MVS/ESA SP5.1 Goal Mode, the sampling is done every 250 milliseconds. The sampling interval is recorded in SMF Type 72 records (R723MTVL).

²Address spaces can have more than one dispatchable unit (that is, more than one TCB or SRB). If an address space did have more than one dispatchable unit, the Workload Manager would accumulate state samples to describe the state of each dispatchable unit. To simplify the discussion, we shall refer to the sample states as being the "state" of the address space, with readers understanding that multiple dispatchable units could simultaneously be waiting on multiple events, or could be using the processor. Thus, for any single sampling, an address space could be counted in more than one state.

³Beginning with OS/390 (Release 3), the Workload Manager calculates execution velocity based on Total Using Samples divided by (Total Using Samples plus Total Delay Samples).

The underlying concept of a sampling process and the analysis of the samples obtained requires that a fairly large number of samples be taken. Basically, sampling normally requires sufficient samples such that the samples can be viewed as representative of the "population" being sampled.

If few address spaces or dispatchable units (TCBs, SRBs, or enclaves) are assigned to a service class, there will be few samples taken. Additionally, if those address spaces or dispatchable units that **are** assigned are primarily idle, few **meaningful** samples will be taken.

Please refer to Section 4 (Chapter 3) for additional information regarding the sampling process, the variables sampled, and how the Workload Manager uses the samples.

CPEXpert examines each SMF Type 72 (Goal) record to determine whether sufficient samples are available for analysis. CPEXpert simply compares the number of meaningful samples with the **MINSAMP** guidance variable contained in USOURCE(WLMGUIDE). CPEXpert retains information for any service class period, by RMF interval, in which the number of meaningful samples is less than the MINSAMP guidance variable.

After all RMF records have been analyzed, CPEXpert determines whether any service class period was inactive (that is, whether few meaningful samples were available for analysis). Either Rule WLM025 or Rule WLM026 may be produced based on this analysis.

CPEXpert produces Rule WLM026 when a service class period at the *highest goal importance* was inactive for more than 75% of the RMF intervals, and when this condition was true for all systems in the sysplex being analyzed.

The following example illustrates the output from Rule WLM026:

RULE WLM026: HIGHEST IMPORTANCE SERVICE CLASS PERIOD HAD FEW SAMPLES

STC_HI (Period 1): This service class period was inactive for over 75% of the RMF measurement intervals being analyzed (inactive means that less than 100 using and delay samples were collected by the SRM for the service class period during the RMF interval). However, the performance goal service class period was given the highest importance. Unfortunately, CPEXpert cannot determine the type of work actually assigned to the service class period. If the work assigned is started tasks, perhaps the started tasks can be assigned to the SYSSTC service class. Please refer to Rule WLM026 in the WLM Component User Manual for a discussion of the reasons for this recommendation.

Suggestion: If only a few meaningful samples are taken related to a service class period with the highest importance, you may wish to consider the following alternatives:

- If the work assigned to the service class period identified by this rule consists of started tasks, you may wish to assign the started tasks to the **SYSSTC** service class. The SYSSTC service class automatically receives a high dispatching priority, and the dispatching priority is above the range of dispatching priorities that are dynamically managed by the Workload Manager⁴.

Since you have indicated that the service class is the *highest* importance, and since few meaningful samples were available for the Workload Manager to analyze, you may find that better performance is achieved if the work is assigned to the SYSSTC service class.

- If you cannot assign the work to the SYSSTC service class, you may wish to combine the work with other important work, and thus both reduce the number of service classes and provide the Workload Manager with more meaningful samples upon which to make its decisions:
 - The Workload Manager makes its decisions mostly based on samples collected from service class periods executing in the sysplex. If few samples are taken for a particular service class period, the Workload Manager cannot validly make decisions regarding the service class period.
 - The Workload Manager will adjust system resource allocation in an attempt to improve performance of only one service class period during a policy adjustment interval⁵.

Adjustment to improve performance of only one service class period is done because the Workload Manager must observe the results of the adjustment: whether the adjustment helped performance, hurt performance, or had no discernable effect.

If adjustments were made to improve performance of more than one service class period, it would be impossible to determine which adjustment helped or hurt which service class period. Consequently,

⁴Address spaces in SYSSTC service class execute at dispatching priority FD (253) if APAR OW19265 is **not** applied, and execute at dispatching priority of FE (254) if OW19265 **is** applied.

⁵The Workload Manager will consider adjustments to improve performance of several service classes (starting with the most important service class which has the highest Performance Index). If performance of the first service class analyzed cannot be appreciably improved, the Workload Manager will select the next worst performing service class at the highest importance, etc. After the Workload Manager has "committed" to a policy adjustment for a service class, it will stop analysis and adjust resources for no other service class.

actions to improve the performance of only one service class period is done during each policy adjustment interval.

The policy adjustment interval is 10 seconds. If too many service class periods have been defined, the Workload Manager may never be able to adjust system resource allocation to help more than a few service class periods (the most important service class periods with the worst performance). Performance of other service class periods may never be improved, or performance improvement actions may take a long time - simply because of the elapsed time necessary for the Workload Manager to make changes, collect data, analyze the effect of the changes, make additional changes, collect more data, and continue the process.

- Perhaps of equal significance is the overhead associated with analyzing service class periods. The Workload Manager tries to improve performance of the service class period with the worst performance at the highest goal importance. Resources may be taken from the least important service class period with the best performance. The Workload Manager will not simply remove and add resources; rather, the Workload Manager will analyze the net value of the planned action.

The Workload Manager will not add resources unless there is an appreciable net gain to the service class period receiving the resources. Within the same goal importance, the Workload Manager will not remove resources from a service class period unless the net gain to the receiver outweighs the net loss to the service class period the resources are being removed from.

The overhead involved with the analysis and decision process increases as the number of service class periods becomes large.

- IBM SRM/WLM developers have indicated that a small number of service class periods is desirable. They have observed that the Workload Manager algorithms typically become increasingly ineffective as the number of service class periods grows. As "rule of thumb" guidance, the developers have stated that most users should define only 25-30 service class periods.

You may wish to assign the work units to a report class, if you are concerned about reporting information related to the work units.

Reference: MVS Planning: Workload Management
MVS/ESA(SP 5): Chapter 8: Defining Service Classes and Performance Goals

OS/390 (V1R1):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V1R2):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V1R3):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V2R4):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V2R5):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V2R6):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V2R7):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V2R8):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V2R9):	Chapter 8: Defining Service Classes and Performance Goals
OS/390 (V2R10):	Chapter 8: Defining Service Classes and Performance Goals
z/OS (V1R1):	Chapter 8: Defining Service Classes and Performance Goals
z/OS (V1R2):	Chapter 8: Defining Service Classes and Performance Goals
z/OS (V1R3):	Chapter 8: Defining Service Classes and Performance Goals
z/OS (V1R4):	Chapter 8: Defining Service Classes and Performance Goals

Thanks: Thanks to Glenn Bowman (Wakefern Food Corporation, NJ) for suggesting this rule.